Coffee-Electrocardiogram Experiment

using latin square design

Young Geun Kim
ygeunkim.github.io

2019711358, Department of Statistics

24 May, 2021

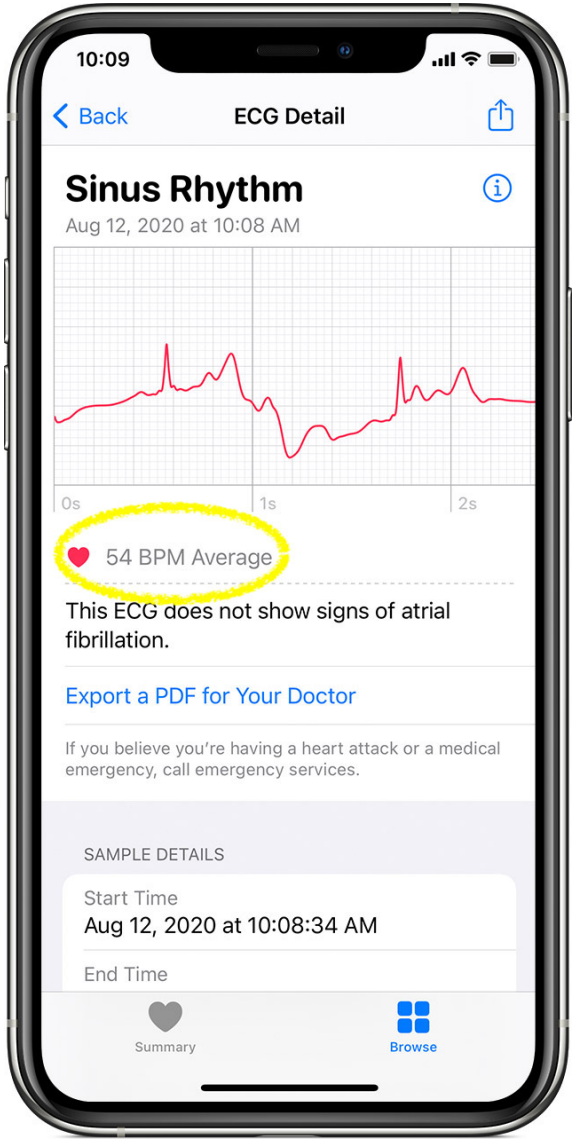
Section 1

Introduction

Electrocardiogram Experiment

Goal

- Does caffeine causally affect ECG or *heart rate*?
- Caffeine: ☕ *coffee*
- Response: ⌚ Average heart rate (BPM)



Latin Square Design

Table 1: Reduced Latin Square

	Column			
Row	1	2	3	4
1	A	B	C	D
2	B	C	D	A
3	C	D	A	B
4	D	A	B	A

- Each treatment once in each row and column
- We allocate 4 treatment levels randomly

Section 2

Design and the Data

Blocking Factors

Caffeine intake depends on the following two factors (한동하, 2018).

Row: Coffee-to-water ratio

1. 1:0 (Espresso, 40 ml)
2. 1:2.5 (Water 100 ml)
3. 1:5 (Water 200 ml)
4. 1:7.5 (Water 300 ml)

Column: Drinking speed

1. <=5 sec
2. 5-15 sec
3. 15-30 sec
4. 30< sec

Figure 1: Coffee (40 ml)

Factor

Intake of caffeine (himynameisabcde, 2020) from Starbucks by Nespresso

1. House blend: 74.5 mg
2. Sumatra: 54.5 mg
3. Decaf espresso roast: 3 mg
4. Just water: 0 mg

Output

Value: Average heart rate

- in BPM
- y_{rc}^{pre}, y_{rc}^{post}: before and after coffee as in Lee (2021)
- Since there exists variation, we consider the difference: $y_{rc} := y_{rc}^{post} - y_{rc}^{pre}$

Measure

- Apple Watch Series 4
- ECG app
- Algorithm version: 1

Randomized Assignment

1. Randomly allocate previous (1, 2, 3, 4) to (A, B, C, D)
2. Assign to above Table 1

```
set.seed(1)
sample(LETTERS[1:4])
#> [1] "A" "C" "D" "B"
```

A. House blend (74.5 mg) - 1
B. Water (0 mg) - 4
C. Sumatra (54.5 mg) - 2
D. Decaf espresso roast (3 mg) - 3

Latin Square

Table 2: Design of the Experiment

	Drinking Speed (in sec.)			
Water	<=5	5-15	15-30	30<
0 ml	HB (74.5)	W (0)	S (54.5)	D (3)
100 ml	W (0)	S (54.5)	D (3)	HB (74.5)
200 ml	S (54.5)	D (3)	HB (74.5)	W (0)
300 ml	D (3)	HB (74.5)	W (0)	HB (74.5)

[1] 'Water' is the coffee-to-water ratio (divide with 40 ml)

[2] Values in the brackets indicate caffeine (in mg)

- Use *reduced latin square*
- Randomization test afterward

Controlling the Other Variables

Coffee

- Drink coffee every morning (between 8:30 a.m. and 9:00 a.m.)
- after eating a piece of bread
- Nespresso machine: Pixie C61 in my home

Measure

- Sitting at the table
- Rest my arms on the table
- Use the same strip
  - Nike sport band
  - of same fit (8-th)
- and other instructions in https://support.apple.com/en-us/HT208955

Section 3

Analysis

Dataset

Table 3: Experiment Data

	Drinking Speed			
water	<=5	5-15	15-30	30<
0 ml	HB, 9	W, 1	S, 7	D, 3
100 ml	W, 2	S, 6	D, 3	HB, 14
200 ml	S, 4	D, 3	HB, 1	W, 0
300 ml	D, 2	HB, 4	W, 2	S, 4

[1] Caffeine: HB > S > D > W

[2] Outputs: after - before taking coffee

Observed Outcomes

Table 4: Observed Outcomes from LS Experiment

	1	2	3	4	Mean
1	9.00	1.0	7.00	3.00	5.00
2	2.00	6.0	3.00	14.00	6.25
3	4.00	3.0	1.00	0.00	2.00
4	2.00	4.0	2.00	4.00	3.00
Mean	4.25	3.5	3.25	5.25	4.06

- Each cell: y_{rc}
- Column: $\bar{y}_{r\cdot}$
- Row: $\bar{y}_{\cdot c}$
- Grand mean: $\bar{y}_{\cdot\cdot}$
- $SSTot = \sum(y_{rc} - \bar{y}_{\cdot\cdot}(\cdot))^2 = 186.938$

Unbiased Estimators

Recall

- Adjusted row effect: $\rho_r = \overline{Y}_{r\cdot}(\cdot) - \overline{Y}_{\cdot\cdot}(\cdot)$
  - unbiased: $\hat{\rho}_r = \bar{y}_{r\cdot}(\cdot) - \bar{y}_{\cdot\cdot}(\cdot)$
- Adjusted column effect: $\gamma_c = \overline{Y}_{\cdot c}(\cdot) - \overline{Y}_{\cdot\cdot}(\cdot)$
  - unbiased: $\hat{\gamma}_c = \bar{y}_{\cdot c}(\cdot) - \bar{y}_{\cdot\cdot}(\cdot)$
- Adjusted treatment effect ()
  - causal estimand
  - $\tau(k) = \overline{Y}_{\cdot\cdot}(k) - \overline{Y}_{\cdot\cdot}(\cdot)$
  - unbiased: $\hat{\tau}(k) = \bar{y}_{\cdot\cdot}(k) - \bar{y}_{\cdot\cdot}(\cdot)$

Causal Estimand

Table 5: Unbiased Estimatior of adjusted treatment effect

House Blend	Sumatra	Decaf	Water
2.9375	1.1875	-1.3125	-2.8125

Caffeine: HB $>$ S $>$ D $>$ W

- Larger caffeine leads to larger effect
- $SSTre = 4\sum \hat{\tau}(k)^2 = 78.688$

Row and Column Effects

Table 6: Unbiased Estimatior of adjusted row effect

100ml	200ml	300ml	400ml
0.9375	2.1875	-2.0625	-1.0625

- Different with what we expected: non-monotonous
- $SSRow = 4\sum \hat{\rho}_r^2 = 44.188$

Table 7: Unbiased Estimatior of adjusted column effect

<=5	5-15	15-30	30<
0.1875	-0.5625	-0.8125	1.1875

- Monotonous, but reverse direction to our original idea
- $SSCol = 4\sum \hat{\gamma}_c^2 = 9.688$
- $SSRow > SSCol$

ANOVA

Table 8: ANOVA Table

Source	Observed			F-Statistic
	DF	Sum Sq	Mean Sq	
water	3	44.19	14.73	1.625
speed	3	9.69	3.23	0.356
coffee	3	78.69	26.23	2.894
Residuals	6	54.38	9.06	
Total	15	186.94		

- ANOVA table from the observed data
- p-value for the treatment: 0.124 (not significant)
- $F_{Tre} = 2.894$ has causal meaning
- Randomize F_{Tre} under sharp null

Sharp Null

Sharp null hypothesis

- of **no effect**
- $H_0 : y_{rc}(1) = y_{rc}(2) = y_{rc}(3) = y_{rc}(4)$
- for each (r, c) pair

Imputing

1. Under the sharp null, impute the missing $Y_{rc}(k)$
2. Apply the formulation

Table 9: Observed Values of the Science Table for the Coffee-ECG Experiment (row 2 and 3: in Appendix)

				Observed $y_{rc}(k)$			
id	water	speed	coffee	HB	W	S	De
Row 1 (Water 0 ml)							
1	1	1	1	9			
2	1	2	4		1		
3	1	3	2			7	
4	1	4	3				3
Row 4 (Water 300 ml)							
13	4	1	3				2
14	4	2	1	4			
15	4	3	4		2		
16	4	4	2			4	

Imputation of Observed Potential Outcomes

Under the sharp null,

Table 10: Imputed Outcomes under the Sharp Null

				Observed $y_{rc}(k)$			
id	water	speed	coffee	HB	W	S	De
Row 1 (Water 0 ml)							
1	1	1	1	9	9	9	9
2	1	2	4	1	1	1	1
3	1	3	2	7	7	7	7
4	1	4	3	3	3	3	3
Row 4 (Water 300 ml)							
13	4	1	3	2	2	2	2
14	4	2	1	4	4	4	4
15	4	3	4	2	2	2	2
16	4	4	2	4	4	4	4

Randomization Test

- Following the same step we learn in the class (Lee, 2021)
- Iterating 2000 times
- p-value is:

$$\text{p-value} = 0.125$$

i.e. Not significant as in ANOVA table

Randomization Distribution

See the histogram.

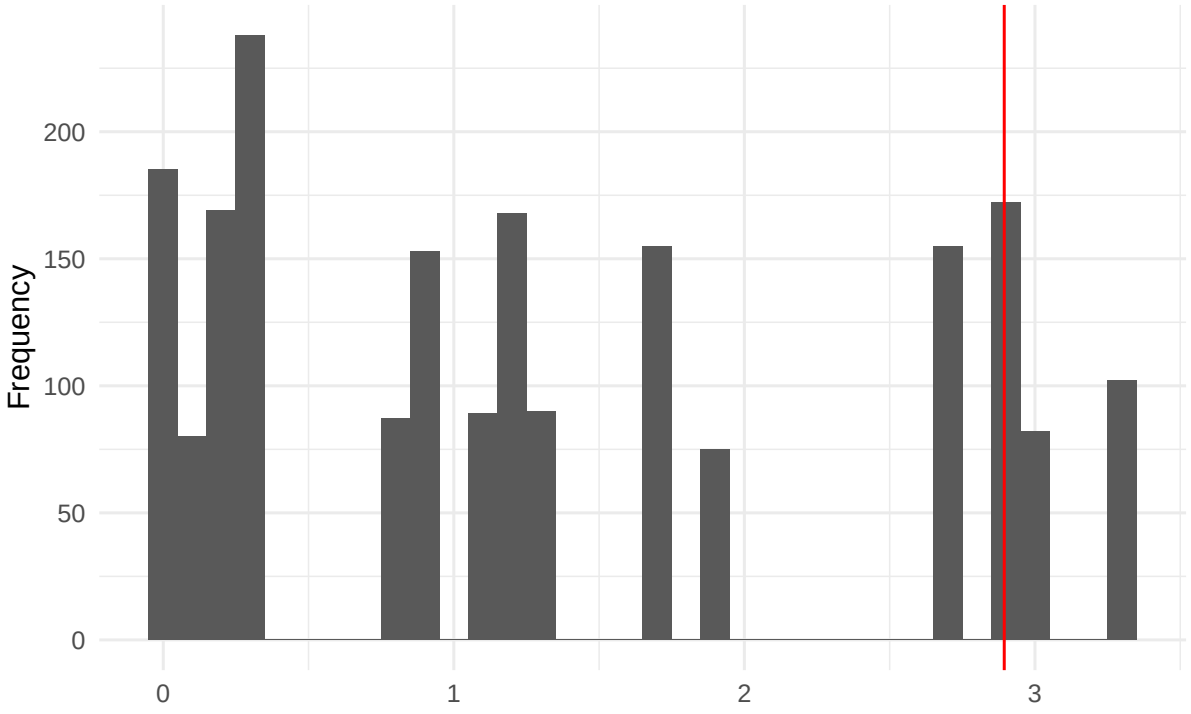


Figure 2: Randomization Distribution of F_{Tre} under the Null

Section 4

Conclusion

Conclusion

- Recall: $H_0 : y_{rc}(1) = y_{rc}(2) = y_{rc}(3) = y_{rc}(4)$
- i.e. Caffeine's effect on average heart rate
- We expected the effect was significant
- However, there was *no significant evidence* (p-value of 0.125) 😭

Discussion: Why this result?

1. Caffeine **tolerance**
   - I have taken coffee everyday
   - Was coffee I have taken too small?
2. Variability: The first treatment (HB) seem too variable

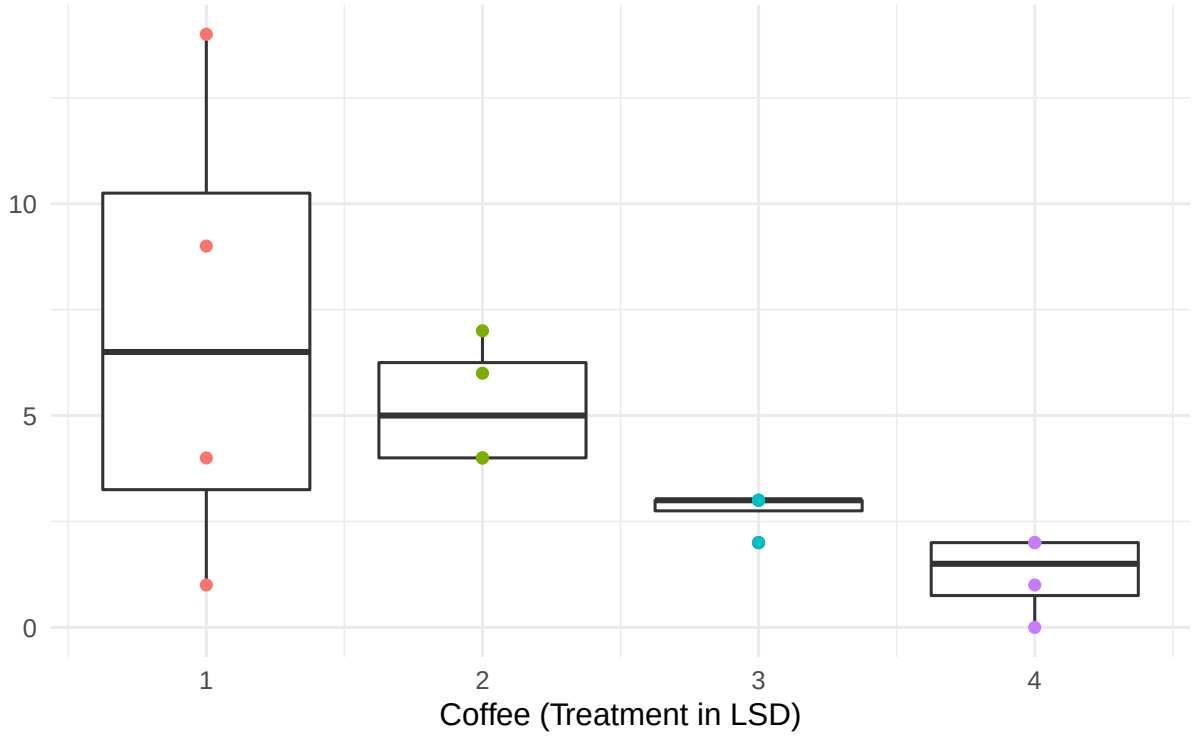


Figure 3: Box Plot for each Treatment

Future study

1. Other samples: Correct arbitrarily set levels
   - Re-define the levels of each factor
2. Re-measure (for Unit 8 and 11)
   - Change $y_{24} = 14$ to 10 and $y_{24} = 1$ to 3

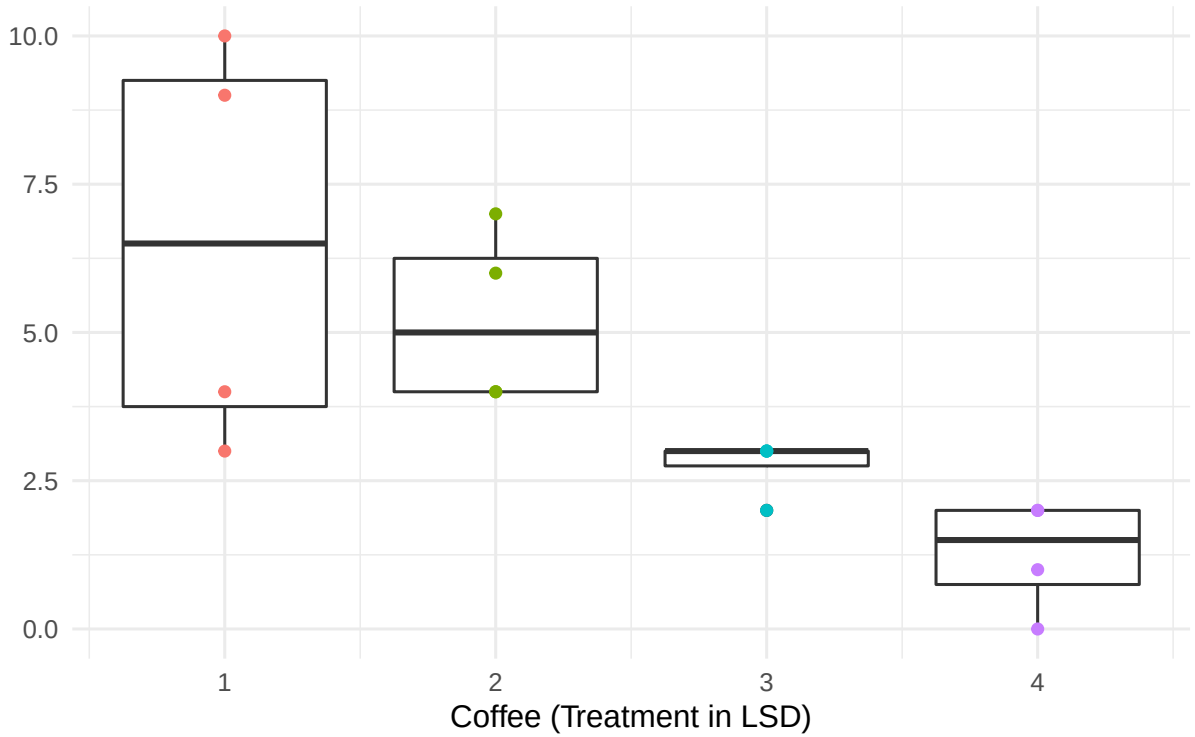


Figure 4: Pseudo-dataset - (2,4) and (3,3) changed

Then 🧐?

- For the pseudo-dataset,
- P-value becomes 0.001 🥳

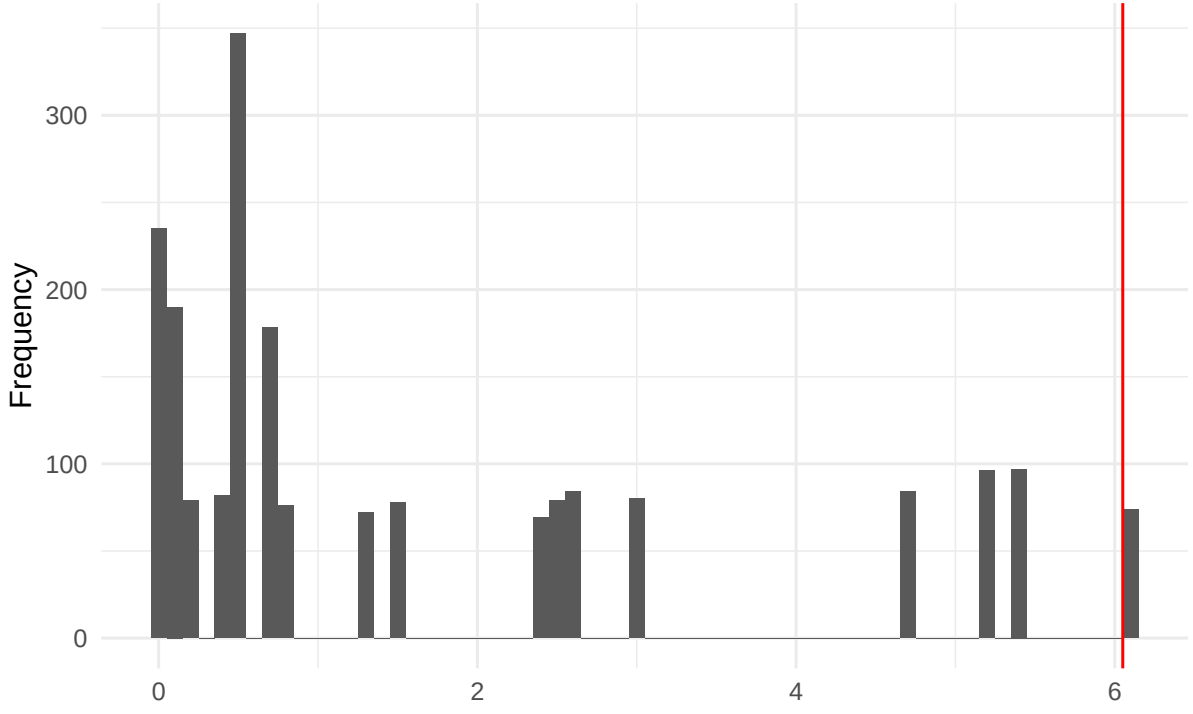


Figure 5: If (2,4) and (3,3) were the different values 🤔

Section 5

Related Contents

Thanks 👋

Feedback 🖐

- Thanks for listening 📊
- Q & A 💡

Codes 💻

- My Github repository: `ygeunkim/ecg-experiment`
  - https://github.com/ygeunkim/ecg-experiment 🖇
  - includes source codes for this analysis 🧰
  - and R markdown files 📄

himynameisabcde (2020). r/nespresso - i recieved the caffiene content numbers for starbucks nespresso pods! https://www.reddit.com/r/nespresso/comments/id31r5/i_recieved_the_caffiene_content_numbers_for/.

Lee, K. (2021). Design and analysis of experiments (sta5031). https://icampus.skku.edu. Accessed: 2021–05-20.

한동하 (2018). [한동하 원장의 웰빙의 역설] 냉커피는 뜨거운 커피와 어떤 차이가 있을까? http://www.k-health.com/news/articleView.html?idxno=37375.